DOI:
10.1016/j.bbr.2016.12.033

Running Title: Spatio-temporal dynamics of outcome evaluation

Computational EEG Modelling of Decision Making Under Ambiguity Reveals Spatio-Temporal Dynamics of Outcome Evaluation

Lee Jollans[1*], Robert Whelan[1*], Louise Venables[2], Oliver H. Turnbull[3], Matteo Cella[2,4] and Simon Dymond† [2,5]

[1] School of Psychology, Trinity College Dublin, Dublin 2, Ireland.

[2] Department of Psychology, Swansea University, Singleton Campus, Swansea SA2 8PP, United Kingdom (s.o.dymond@swansea.ac.uk).

[3] School of Psychology, Bangor University, Bangor, Gwynedd, LL57 2DG, United Kingdom.

[4] Department of Psychology, Institute of Psychiatry, Psychology & Neuroscience, King's College London, De Crespigny Park, SE5 8AF, London, United Kingdom.

[5] Department of Psychology, Reykjavík University, Menntavegur 1, Nauthólsvík, 101 Reykjavík, Iceland.

† Corresponding author

*Equal contribution

Abstract

Complex human cognition, such as decision-making under ambiguity, is reflected in dynamic spatio-temporal activity in the brain. Here, we combined event-related potentials (ERPs), with their high temporal resolution and moderate spatial resolution afforded by high-density arrays, with computational modelling of the time course of decision-making and outcome evaluation. We contrasted four computational models (Expectancy-Valence, Prospect Valence Learning (PVL), PVL-Delta, and Value Plus Perseverance (VPP) model) and found that the VPP model provided the best post hoc fit. Measures of choice probability generated using the VPP model, as well as objective trial outcomes, were applied as regressors in a general linear model of the EEG signal to create a three-dimensional spatio-temporal characterization of task-related neural dynamics. We observed that outcome valence, outcome magnitude, and VPP choice probability are expressed in distinctly separate components of the ERP. Our findings show, for the first time, model-based analysis of the spatio-temporal dynamics of outcome evaluation in complex human decision-making.

Keywords: computational models, decision making, EEG, model based.

A major challenge in cognitive neuroscience research on decision-making under ambiguity is delineating how prediction errors and task outcomes guide future behavior (Frank et al. 2004; Redish et al. 2007; Gläscher et al. 2010; Cockburn et al. 2014). Tasks like the Iowa Gambling Task (IGT) are a popular measure of decision making under ambiguity (e.g., Brevers et al. 2013; Xiao et al. 2013; Halfmann et al. 2014). Advantageous performance on the IGT is based on approximations of long-term consequences rather than exact calculations (Christakou et al. 2009), and choice behavior typically shifts across trials as participants learn to make more advantageous selections with increasing knowledge of the outcome contingencies (Gansler et al. 2011).

Computational models such as the Expectancy-Valence (EV) model (Busemeyer and Stout 2002), the Perseverance Valence Learning (PVL) model, the PVL-Delta model (Ahn et al. 2008, 2011; Fridberg et al. 2010), and the Value Plus Perseverance (VPP) model (Worthy et al. 2013), consider factors such as the attention given to outcome valence (i.e., to wins vs. losses), how the recency of feedback affects future decisions, and how choices are influenced by experience (i.e., to what extent choices are random). Such models assume that the valence experienced on each trial informs a probabilistic choice mechanism, and quantifies outcome expectation and prediction error on an individual trial-by-trial basis, thereby estimating individuals' subjective experiences of the task and task-expectations, rather than objective task outcomes (Yechiam and Busemeyer 2005). The approximation of the experienced valence on each trial is calculated using a utility function that accounts for the relative value placed on positive or negative outcomes. The experienced valence value is then used to calculate, using a learning rule that accounts for choice recency, the *expected* valence for each deck on the following trial. Uniquely, the VPP model also includes a measure of perseverance that takes into account the likelihood of selecting the same deck again based on outcome valence. The probability that each of the four decks will be selected on the next trial is calculated using a sensitivity function, which includes a parameter estimating the participant's consistency in choice responding. Generally, the PVL-Delta model has been found to generate choice patterns more accurately than the EV or PVL models (Steingroever et al. 2014), while the VPP model provides the best post-hoc fit to observed choice patterns when compared to the EV, PVL, and PVL-Delta models (Steingroever et al. 2016; Ahn et al. 2014). Parameter values from these computational models have revealed differences in behavioral decision-making between control subjects and clinical populations (Cella et al. 2010).

Here, we used event-related potentials (ERPs), with their high temporal resolution and moderate spatial resolution afforded by high-density arrays, to examine the time course of decision-making and outcome evaluation with their temporally proximate and often overlapping stages. ERPs associated with reward anticipation and feedback processing are sensitive to a number of parameters, such as the valence, magnitude and likelihood of the outcome (Holroyd et al. 2004, 2011; Hajcak et al. 2005; Wu and Zhou 2009; Talmi et al. 2012; Fuentemilla et al. 2013). However, most ERP studies of feedback learning have contrasted average ERPs over different trial types rather than considering the trial-by-trial progression of outcome processing (e.g., Hajcak et al. 2006), although there been recent investigations of trial-by-trial modeling using EEG (Larsen and O'Doherty 2014).

The present study reveals the spatio-temporal neural dynamics of decision-making under ambiguity using approximations of outcome expectations generated using a best-fitting computational model of IGT performance. We first gathered behavioral and EEG data from participants while they performed the IGT. We then identified which of four computational models provided the best fit to the performance of all participants. Next, we used the model trial-by-trial component values to form regressors in a general linear model of the signal. Our objective was to identify, for the first time, subjective components of decision making under ambiguity, and to identify if and how their neural signature differs from that of absolute task outcomes and trial characteristics, thus providing insight into the nature and time course of the interplay between key cognitive processes and task experience when making decisions.

Materials and Methods

Participants

Twenty healthy, right-handed adults (9 female), 19 to 38 years old (*mean age* = 24.9 years, *SD* = 4.8 years) participated and were reimbursed with £10 that was not contingent on performance. The Department of Psychology Ethics Committee, Swansea University, approved all procedures.

Iowa Gambling Task

We used a computerized variant of the original IGT (Bechara et al. 1994) in which participants were instructed to select cards from four concurrently available decks (labeled A, B, C and D). Deck locations were randomly varied across participants. Trials were preceded by a 2s choice appraisal interval, during which choices could not be made, as the four individual decks and the text, “Please consider your choice” appeared on screen. After this, choices were made using the mouse (the cursor was centered at the start of every trial). An initial ‘loan’ of £1000 virtual money, displayed at the bottom of the screen, was updated immediately following choices accompanied by text stating the amount of money gained and/or lost. Decks A and B (termed ‘disadvantageous’) resulted in long-term loss (£250 loss per 10 trials), whereas decks C and D (termed ‘advantageous’) resulted in long-term gain (£250 gain per 10 trials). Participants always won £100 if they selected a card from the advantageous decks, and £50 if they selected a card from the disadvantageous decks. Losses varied between £150 and £350 for deck A; £1250 for deck B; £25 to £75 for deck C; and

£250 for deck D. Decks A and C resulted in frequent losses (on 50% of trials), whereas decks B and D resulted in infrequent losses (on 10% of trials). Onscreen feedback was displayed for 10 s, before a 2 s inter-trial interval. The task ended after 100 trials. After every block of 20 choices, subjective awareness ratings were made of the relative ‘‘goodness’’ or ‘‘badness’’ of each deck (Bowman et al. 2005; Cella et al. 2007) using a slider-scale from 0 (‘*very bad*’) to 10 (‘*very good’*).

Analysis 1: Computational modeling of IGT performance

The deck chosen on trial t is denoted $D(t)$. The reward received on each trial is denoted $R(t)$, and the loss on each trial is denoted $L(t)$, such that if deck D_3 (a disadvantageous deck) were chosen on trial $t = 9$ (i.e. $D(9) = D_3$) then $R(D(9))$ = £100 and $L(D(9))$ = £1250. The absolute monetary outcome on each trial is denoted $X(t)$.

The EV model, PVL model, PVL-Delta model, and VPP model were compared. These models all follow the same shape: an approximation of the experienced valence $u(t)$ on each trial t is calculated using a utility function, based on $R(t)$ and $L(t)$. The utility functions take into account the relative value placed on positive or negative outcomes. The experienced valence value is then used to calculate the expected valence $Ev(t+1)_j$ for the selected deck j on the following trial. Ev is calculated using a learning rule, which includes a parameter representing the effect of recency on choice behavior. Uniquely, the VPP model also includes a measure of perseverance which takes into account the likelihood of selecting the same deck again based on outcome valence. Finally, the probability $Pr[D(t+1)=j]$ that each of the four decks will be selected on the next trial is calculated using a sensitivity function, which includes a parameter estimating the consistency with which participants respond, and a Softmax action-selection rule. The functions used in each model are presented in Table 1.

Insert Table 1 About Here

Calculating model fit

A custom Matlab script was prepared (Supplementary Materials), which calculated model fit for each participant for each of the four models. For each parameter three starting parameter values (the minimum and maximum parameter values as defined by the models, and an intermediate value) were entered into the Matlab *fminunc* function, which maximized the log likelihood value for each parameter combination using a quasi-newton algorithm and 5000 function evaluations. This resulted in between 27 (for the EV model) and 6561 (for the VPP model) parameter combinations being tested. The parameter combination which resulted in the highest chi value from a likelihood ratio test was taken as the best model fit for each participant.

Results: Behavioral and Computational Modelling

The mean number of times advantageous and disadvantageous decks, as well as frequent and infrequent loss decks were selected per block of 20 trials was calculated for each participant. A two-way non-repeated ANOVA found a significant interaction effect of task block and advantageous or disadvantageous deck on task choices (F = 8.33, df = 4, p < .001, η_p^2=.149; see Figure 1). Simple main effects analyses showed that disadvantageous decks were more frequently selected than advantageous decks during the first task block (M_1 = 13, SD_1 = 2.43, p_1 < .001) and the second task block (M_2 = 11.35, SD_2 = 2.70, p_2 = .017). During the last task block, the number of times advantageous decks were selected was significantly higher than the number of times disadvantageous decks were selected (M = 11.25, SD = 3.76, p = .027).

Insert Figure 1 About Here

A two-way non-repeated ANOVA found no significant interaction effect of task block and frequent or infrequent loss deck on task choices (F = 1.51, df = 4, p =.202, η_p^2=.031; see

Figure 2). While the main effect of block was not significant (F = 0, df = 4, p =1, η_p^2=0), there was a significant main effect of deck type (F = 82.86, df = 1, p <.001, η_p^2=.304), with infrequent loss decks being selected significantly more frequently that frequent loss decks (M=12.03, SD=3.14).

Insert Figure 2 About Here

A one-way non-repeated ANOVA of awareness ratings found a significant effect of task block (F = 2.59, df = 4, p = .042, η_p^2=.103; see Figure 1). The Tukey post hoc criterion revealed no significant differences, but the contrast between net ratings during the first (M = .184) and second task block (M = 2.263, p = .061) and the first and third task block (M = 2.316, p = .052) approached significance, indicating that awareness of deck contingencies during the first task block was lower than during the subsequent blocks. Table 2 displays the mean correct deck prediction percentages (see Supplementary Table 1 for the model with the highest correct deck predictions per subject).

Insert Table 2 About Here

Behavioral data and the VPP model

The percentage of total correct choice prediction was higher for the PVL model than the VPP model. However, the VPP model provided the best model fit for the majority of participants (see Table 3). Furthermore, a series of logistic regressions using the choice probability values calculated for each deck by the computational models to predict whether each deck was chosen or not revealed that the VPP model generated the most accurate choice prediction (see Figure 3). The distribution of the optimal values for each parameter from the VPP model across all participants is reported in Supplementary Table 2. Deck choices were significantly positively associated with the decks with the highest choice probability across all trials (r=0.28, p=2x10^{-36}). Choice probability was significantly positively associated with subjective awareness ratings across 5 task blocks for disadvantageous decks (r=0.29, p=0.003) but not for advantageous decks (r=0.03, p=0.75). Choice probability was significantly positively associated with the number of times each deck was chosen across 5 task blocks for advantageous decks (r=0.75, p=2x10^{-37}), for disadvantageous decks (r=0.83, p=7x10^{-53}), for decks with infrequent losses (r=0.79, p=2x10^{-44}), and for decks with frequent losses (r=0.74, p=4x10^{-36}).

Insert Table 3 About Here

Insert Figure 3 About Here

Analysis 2: ERP correlates of IGT performance

EEG recording

EEG data were recorded in a soundproofed room using the ActiveTwo Biosemi™ electrode system from 134 electrodes (128 scalp electrodes) organized according to the 10-5 system (Oostenveld and Praamstra 2001), digitized at 512 Hz.

EEG analysis

EEG preprocessing and artifact rejection was performed using the Fully Automated Statistical Thresholding for EEG artifact Rejection toolbox (FASTER; http://sourceforge.net/projects/faster; Nolan et al. 2010), implemented in EEGLAB (Delorme and Makeig 2004) under Matlab 7.12. EEG data were filtered (1–95 Hz, with a notch filter at 50 Hz). Epoch length was initially set to -3 s to 2 s for the choice appraisal interval (marker set to onset of appraisal interval) and the outcome evaluation phase (marker set to onset of outcome). EEG data from one participant was excluded due to poor data quality.

EEG data were processed in SPM8 (http://www.fil.ion.ucl.ac.uk/spm). Data from each participant were transformed into two-dimensional sensor-space (interpolated from the

128 scalp channels), over peri-stimulus times from -100–600 ms for the feedback processing phase, thus producing a three-dimensional spatio-temporal characterization of the ERP. Baseline was corrected from 100 ms before cue presentation. The EEG timeseries data were subsequently parcellated based on both spatial and temporal domains. Data were averaged in 64 spatial bins, and across time segments of 25.4 ms (resulting in 23 time bins in the outcome phase).

Outcome measures

For each participant, three variables were used as regressors in a GLM with the parcellated outcome data from the same trials: the valence and the magnitude of the outcome (objective outcome measures), and the trial-by-trial choice probability for the selected deck calculated using the VPP model. The temporal and spatial properties of associations between regressors and the EEG timecourse across the whole outcome interval were examined. Associations between valence, magnitude and choice probability and two ERP components that consistently occur following feedback, the FRN and the P3, were examined.

Significance testing

A linear regression was carried out for each regressor individually. This resulted in a beta weight being generated for each regressor and each bin. The same calculations were also carried out using a random permutation of the model regressors (i.e. the values of each regressor were shuffled), which resulted in a baseline, or ‘null’ distribution. For each regressor and each of the bins a one-sample t-test was carried out using the beta values for each participant, as well as the beta values from the random label permutations. For each regressor the bins in which the test statistic was larger than the 95th percentile of the distribution of test statistic values for the beta weights generated using random label permutations were deemed significantly associated with the regressor.

Results: Spatio-temporal dynamics

Associations between outcome measures and the EEG timecourse

Of the 1472 (64 spatial by 23 temporal) bins, 312 bins (21%) showed a significant association with at least one of the regressors, with almost all (289 bins, 20%) uniquely associated with one regressor. Twenty bins (1.4%) were associated with two regressors, and 3 bins (0.2%) were associated with all three regressors. The number of bins each regressor was significantly associated with is presented in Table 4.

Insert Table 4 About Here

Bins associated with objective task outcomes. ERPs associated with outcome magnitude occurred most strongly in the first 25-76ms after feedback presentation (see Figure 4B). Magnitude was also significantly associated with the ERP between 229ms and 482ms after feedback.

Insert Figure 4 About Here

Valence was significantly associated with the ERP throughout most of the outcome processing interval, up to about 500 ms after feedback presentation (see Figure 4B). The largest number of significant associations between the ERP and valence occurred between 279ms and 457ms after feedback presentation, although there were also a large number of significant associations between 152ms and 178ms after feedback presentation. The ERP in 13 bins was significantly associated with both magnitude and valence, with the majority of these associations occurring between 279ms and 482ms after feedback presentation (Supplementary Table 3).

Bins associated with VPP choice probability. The VPP model choice probability was associated with the ERP from 127ms to 558ms after feedback presentation (see Figure 4B).

The largest number of significant associations occurred in the interval from 254ms to 305ms after feedback presentation, with an additional elevation in the number of significant association toward the end of the outcome processing interval.

Bins associated with objective variables and VPP choice probability. Choice probability and magnitude, valence, or both magnitude and valence were significantly associated with the ERP in 10 bins across five temporal intervals. These associations occurred predominantly between 279 and 381ms after feedback presentation (Supplementary Table 4).

Associations between outcome measures and predefined ERPs

Feedback related Negativity (FRN). Based on the observed EEG signal (Figure 4A) the FRN was defined as the interval between 178ms and 355ms. During the FRN time interval, choice probability was associated with 39 bins, valence was associated with 83 bins, and outcome magnitude was associated with 9 bins (Figure 5A).

P3. Based on the observed EEG signal (Figure 4A), the P3 was defined as the interval between 355ms and 482ms. During the P3 time interval, choice probability was associated with 12 bins, valence was associated with 83 bins, and outcome magnitude was associated with 15 bins (Figure 5B).

Insert Figure 5 About Here

Discussion

We used an established computational model of decision making during the IGT to reveal the electrophysiological correlates of feedback processing in a spatio-temporal characterization of the ERP. A comparison of four models (EV, PVL, PVL-Delta, and VPP) revealed that the VPP model demonstrated the best post-hoc fit, in line with previous findings

(Steingroever et al. 2016; Ahn et al. 2014). The choice probability value created using the VPP model mapped closely onto choice frequency for all decks, and was correlated with subjective awareness ratings for disadvantageous decks. The trial-by-trial values for choice probability calculated by the VPP model were subsequently used as regressors in a GLM of the EEG timecourse during the feedback processing interval alongside outcome valence and outcome magnitude. This revealed a strong effect of the VPP choice probability variable on the EEG time-course during the FRN time interval.

Across the entire outcome interval, valence was significantly associated with more than twice as many spatiotemporal bins of the ERP as the other model regressors. The VPP model choice probability variable was uniquely associated with nearly 5% of all spatiotemporal bins. Given the large number of bins which were significantly associated with one of the three regressors (~20%), it is surprising that fewer than 1 in 10 of these bins were associated with more than one regressor. This indicates that the three elements of feedback processing evaluated here (outcome valence, outcome magnitude, and subjective outcome expectations) show distinctly different temporal properties in terms of their expression in the ERP.

Based on the observed EEG timecourse, the FRN ERP occurred between approximately 180ms and 360ms after feedback presentation. This is a longer interval than is often used to examine the FRN (Weinberg et al. 2014; Cui et al. 2013; Fuentemilla et al. 2013). However, the regressor which was expressed most strongly in the EEG in the interval more traditionally associated with the FRN (between about 230ms to 300ms after feedback presentation) was the VPP model choice probability, which is consistent with the role in prediction error attributed to the FRN. The FRN has also been associated with the likelihood of an outcome (Holroyd et al. 2004; Fuentemilla et al. 2013), which provides conceptual support for the VPP model as a valid measure of participants' choice certainty. Considering that the VPP model choice probability variable is thought to be a reflection of the decision-makers' subjective valuation of different possible outcomes, our findings may also support the hypothesis brought forward by Talmi and colleagues, that the FRN reflects prediction errors in terms of outcome saliency rather than reward prediction errors (Talmi et al. 2012, 2013).

Past research has also found that the FRN is associated with outcome valence (Yeung and Sanfey 2004; Hajcak et al. 2006), whereas the P3 appears not to be associated with valence (Yeung and Sanfey 2004). This clear dissociation between the FRN and P3 is inconsistent with our findings relating to the expression of valence in the ERP. Valence was strongly associated with the ERP between 280ms and 460ms after feedback presentation, which includes both a late component of the FRN, as well as almost the entire P3 ERP. Outcome magnitude showed only a small number of significant associations with the EEG timecourse across the entire outcome interval. While there were some significant association between outcome magnitude and both the FRN and the P3, the quantity of these associations was quite small. While previous research suggests that the P3 is sensitive to magnitude (Yeung and Sanfey 2004), there have been conflicting findings with regard to the effect of magnitude on the FRN. A recent meta-analysis suggested that the FRN does show a strong main effect of reward magnitude (Sambrook and Goslin 2015), while other studies suggest that the FRN is not associated with magnitude (Yeung and Sanfey 2004; Hajcak et al. 2006; Cui et al. 2013).

From about 500ms after feedback presentation choice probability was the only variable which was consistently associated with the ERP in a number of scalp locations. This is similar to findings by (Talmi et al. 2012), who also observed that outcome probability, but not outcome valence, was expressed in the MEG signal 500ms after outcome presentation. Since the objective elements of task outcome do not find a strong expression in the ERP anymore at this point it can be concluded that this reflects a late deliberation component in outcome processing.

In conclusion, the present study used a well validated model of choice behaviour in the IGT to map the spatiotemporal expression of subjective choice certainty in the ERP during outcome processing. This revealed that participants' subjective choice valuations were strongly associated with the FRN, providing a novel perspective on this well researched ERP component. Furthermore, findings suggest that the temporal progression of outcome processing during the FRN time window may warrant further investigation with respect to when valence and judgements of the likelihood of an outcome are processed. The degree to which objective measures of trial feedback such as outcome valence and magnitude are reflected in the well-established P3 component also warrants further investigation.

Funding

This work was supported by the Wales Institute of Cognitive Neuroscience.

Acknowledgements

LJ is supported by the Irish Research Council (IRC) under project no. GOIPG/2014/418.

References

Ahn WY, Busemeyer JR, Wagenmakers EJ, Stout JC. 2008. Comparison of decision learning models using the generalization criterion method. Cogn Sci. 32:1376-402.

Ahn WY, Krawitz A, Kim W, Busemeyer JR. 2011. A model-based fMRI analysis with hierarchical Bayesian parameter estimation. J Neurosci Psychol Econ. 4:95-110.

Ahn WY, Vasilev G, Lee SH, Busemeyer JR, Kruschke JK, Bechara A, Vassileva J. 2014. Decision-making in stimulant and opiate addicts in protracted abstinence: evidence from computational modeling with pure users. Front Psychol. 5:849.

Bechara A, Damasio A, Damasio H, Anderson S. 1994. Insensitivity to future consequences following damage to human prefrontal cortex. Cognition. 50:7-15.

Bowman C, Evans C, Turnbull O. 2005. Artificial time constraints on the Iowa Gambling Task: The effects on behavioural performance and subjective experience. Brain Cogn. 57:21-25.

Brevers D, Bechara A, Cleeremans A, Noël X. 2013. Iowa Gambling Task (IGT): twenty years after–gambling disorder and IGT. Front Psychol. 4:665.

Busemeyer JR, Stout JC. 2002. A contribution of cognitive decision models to clinical assessment: decomposing performance on the Bechara gambling task. Psychol Assess. 14:253-262.

Cella M, Dymond S, Cooper A. 2010. Impaired flexible decision-making in major depressive disorder. J Affect Disord. 124:207–210.

Cella M, Dymond S, Cooper A, Turnbull O. 2007. Effects of decision-phase time constraints on emotion-based learning in the Iowa Gambling Task. Brain Cogn. 64:164-169.

Christakou A, Brammer M, Giampietro V, Rubia K. 2009. Right ventromedial and dorsolateral prefrontal cortices mediate adaptive decisions under ambiguity by integrating choice utility and outcome evaluation. J Neurosci. 29:11020–11028.

Cockburn J, Collins A, Frank M. 2014. A reinforcement learning mechanism responsible for the valuation of free choice. Neuron. 83:551-557.

Cui J, Chen Y, Wang Y, Shum D, Chan R. 2013. Neural correlates of uncertain decision making: ERP evidence from the Iowa Gambling Task. Front Hum Neurosci. 7:776.

Delorme A, Makeig S. 2004. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. J Neurosci Methods. 134:9-21.

Frank M, Seeberger L, O'Reilly R. 2004. By carrot or by stick: Cognitive reinforcement learning in Parkinsonism. Science. 306:1940–1943.

Fridberg DJ, Queller S, Ahn W-YY, Kim W, Bishara AJ, Busemeyer JR, Porrino L, Stout JC. 2010. Cognitive mechanisms underlying risky decision-making in chronic cannabis users. J Math Psychol. 54:28–38.

Fuentemilla L, Cucurell D, Marco-Pallarés J, Guitart-Masip M, Morís J, Rodríguez-Fornells A. 2013. Electrophysiological correlates of anticipating improbable but desired events. Neuroimage. 78:135–144.

Gansler D, Jerram M, Vannorsdall T, Schretlen D. 2011. Comparing alternative metrics to assess performance on the Iowa Gambling Task. J Clin Exp Neuropsychol. 33:1040-1048.

Gläscher J, Daw N, Dayan P, O'Doherty J. 2010. States versus rewards: Dissociable neural prediction error signals underlying model-based and model-free reinforcement learning. Neuron. 66:585–595.

Hajcak G, Holroyd C, Moser J, Simons R. 2005. Brain potentials associated with expected and unexpected good and bad outcomes. Psychophysiology. 42:161–170.

Hajcak G, Moser J, Holroyd C, Simons R. 2006. The feedback-related negativity reflects the binary evaluation of good versus bad outcomes. Biol Psychol. 71:148-54.

Halfmann K, Hedgcock W, Bechara A. 2014. Functional neuroimaging of the Iowa Gambling Task in older adults. Neuropsychology. 28:870

Holroyd C, Krigolson O, Lee S. 2011. Reward positivity elicited by predictive cues. Neuroreport. 22:249-252.

Holroyd C, Larsen J, Cohen J. 2004. Context dependence of the event-related brain potential associated with reward and punishment. Psychophysiology. 41:245–253.

Larsen T, O'Doherty JP. 2014. Uncovering the spatio-temporal dynamics of value-based decision-making in the human brain: a combined fMRI-EEG study. Philos Trans R Soc Lond, B, Biol Sci. 369:20130473.

Nolan, Whelan, Reilly RB. 2010. FASTER: Fully Automated Statistical Thresholding for EEG artifact Rejection. J Neurosci Methods. 192:152-62.

Oostenveld R, Praamstra P. 2001. The five percent electrode system for high-resolution EEG and ERP measurements. Clinical Neurophysiol. 112:713-9.

Redish A, Jensen S, Johnson A, Kurth-Nelson Z. 2007. Reconciling reinforcement learning models with behavioral extinction and renewal: Implications for addiction, relapse, and problem gambling. Psychol Rev. 114:784-805.

Sambrook TD, Goslin J. 2015. A neural reward prediction error revealed by a meta-analysis of ERPs using great grand averages. Psychol Bull. 141:213–235.

Steingroever H, Wetzels R, Wagenmakers EJ. 2014. Absolute performance of reinforcement-learning models for the Iowa Gambling Task. Decision. 1:161.

Steingroever H, Wetzels R, Wagenmakers EJ. 2016. Bayes factors for reinforcement-learning models of the Iowa gambling task. Decision. 3:115.

Talmi D, Atkinson R, El-Deredy W. 2013. The Feedback-Related Negativity Signals Salience Prediction Errors, Not Reward Prediction Errors. J Neurosci. 33:8264–8269.

Talmi D, Fuentemilla L, Litvak V, Duzel E, Dolan R. 2012. An MEG signature corresponding to an axiomatic model of reward prediction error. NeuroImage. 59:635–645.

Weinberg A, Riesel A, Proudfit G. 2014. Show me the money: the impact of actual rewards and losses on the feedback negativity. Brain Cogn. 87:134–139.

Worthy D, Pang B, Byrne K. 2013. Decomposing the roles of perseveration and expected value representation in models of the Iowa gambling task. Front Psychol. 4:640.

Wu Y, Zhou X. 2009. The P300 and reward valence, magnitude, and expectancy in outcome evaluation. Brain Res. 1286:114–122.

Xiao L, Wood S, Denburg N, Moreno G, Hernandez M, Bechara A. 2013. Is there a recovery of decision-making function after frontal lobe damage? A study using alternative versions of the Iowa Gambling Task. J Clin Exp Neuropsychol. 35:518–529.

Yechiam E, Busemeyer JR. 2005. Comparison of basic assumptions embedded in learning models for experience-based decision making. Psychon Bull Rev. 12(3):387-402.

Yeung N, Sanfey A. 2004. Independent coding of reward magnitude and valence in the human brain. J Neurosci. 24:6258–6264.

Tables

Table 1. Model formulae for the EV, PVL, PVL-Delta, and VPP models.

Concept	Model(s)	Model equation	Free parameters (range)
Utility function	EV	$u(t) = (1 - w) * R(t) + w * L(t)$	w: Attention weight (0,1)
	PVL, PVL-Delta, VPP	$u(t) = \begin{cases} X(t)^{\alpha} \; if \; X(t) \geq 0 \\ = \lambda * \lvert X(t) \rvert^{\alpha} \; if \; X(t) < 0 \end{cases}$	α: Shape (0,1) λ: Loss aversion (0,5)
Learning rule	EV, PVL-Delta, VPP	$Ev_j(t+1) = Ev_j(t) + \phi * (u(t) - Ev_j(t))$	ϕ: Recency (0,1)
	PVL	$Ev_j(t+1) = A * Ev_j(t) + \delta_j(t) * u(t)$	A: Decay (0,1) $\delta_j = \begin{cases} 1 \; if \; Deck \; j \; was \; chosen \\ 0 \; if \; Deck \; j \; was \; not \; chosen \end{cases}$
Perseveration function	VPP	Chosen deck: $P_j(t+1) = \begin{cases} k * P_j(t) + \varepsilon_{pos} \; if \; X(t) \geq 0 \\ k * P_j(t) + \varepsilon_{neg} \; if \; X(t) < 0 \end{cases}$ Unchosen decks: $P_j(t+1) = k * P_j(t)$ $V_j(t+1) = \omega * Ev_j(t+1) + (1 - \omega) * P_j(t+1)$	k: Decay (0,1) ε_{pos}: Impact of gain on perseverance (-1,1) ε_{neg}: Impact of loss on perseverance (-1,1) ω: Reinforcement learning (0,1)
Sensitivity function	EV	$\theta(t) = \left(\frac{t}{10}\right)^{c}$	c: Consistency (-2,2)
	PVL, PVL-Delta, VPP	$\theta(t) = 3^{c} - 1$	c: Consistency (0,5)
Softmax action-selection rule	EV, PVL, PVL-Delta	$Pr[D(t+1) = j] = \frac{e^{\theta(t) Ev_j}}{\sum_{k=1}^{4} e^{\theta(t) Ev_k}}$	
	VPP	$Pr[D(t+1) = j] = \frac{e^{\theta(t) V_j}}{\sum_{k=1}^{4} e^{\theta(t) V_k}}$	

Table 2. Mean percentage of correct deck predictions for each model across participants.

	Correct deck (%)	Advantageous vs. disadvantageous deck (%)	Frequent vs. infrequent loss deck (%)
EV	27.72	40.55	43.88
PVL	59.34	72.47	72.12
PVL-Delta	42.07	63.58	59.69
VPP	43.74	63.38	62.68

Table 3. Number of participants (out of 20) that each model provided the best fit for, using different metrics.

	Akaike Information Criterion	**Bayesian Information Criterion**
EV	1	1
PVL	2	6
PVL-Delta	3	4
VPP	14	9

Table 4. Percentage of all 1472 bins that were significantly associated with each regressor.

Regressor	**Bins associated with the regressor (%)**	**Bins uniquely associated with regressor (%)**
Outcome magnitude	3.74	2.51
Outcome valence	13.72	12.29
VPP Choice Probability	5.50	4.82

Captions to Figures

Figure 1. Deck choice frequency, subjective awareness ratings, and VPP model choice probability for advantageous and disadvantageous decks.

Figure 2. Deck choice frequency and VPP model choice probability for frequent and infrequent loss decks.

Figure 3. Receiver Operating Characteristic (ROC) curve for the logistic regression using the choice probability values calculated for each deck by the VPP model to predict whether each deck was chosen or not. The area under the curve (AUC) for the VPP model (AUC=.73) exceeded that for the PVL model (AUC=.71), for the EV model (AUC=.60), and for the PVL Delta model (AUC=.53).

Figure 4. (A) ERPs in each temporal bin from anterior to posterior, averaged over left and right. (B) Percentage of spatial bins from left to right in which each regressor (outcome magnitude, valence, and choice probability) was significantly associated with the ERP in each time bin at each of eight spatial locations from anterior to posterior.

Figure 5. (A) Feedback Related Negativity (FRN) ERP component with activation averaged across the FRN timecourse (178ms after feedback to 355ms after feedback), with percentage of time bins during the FRN time interval for which each regressor (outcome magnitude, valence, and choice probability) was significantly associated with the ERP. (B) P300 ERP component with activation averaged across the P3 timecourse (355ms after feedback to

482ms after feedback), with percentage of time bins during the P3 time interval for which each regressor (outcome magnitude, valence, and choice probability) was significantly associated with the ERP.